```
OM of: US-09-126-945B-2_COPY_236_243 to: N_Geneseq_0601:*  out_format : pfs

Date: Sep 7, 2001  6:30 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09126945/runat_07092001_152109_3686/app_query.fasta_1.487
 -DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=olig.rng -GAPOP=4.500
 -GAPEXT=0.050 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=60.000 -XGAPEXT=60.000
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=oligo
 -TRANS=human40.cd1 -LIST=45 -DOCALIGN=200 -THR_SCORE=quality
 -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09126945_@CGN1_1_300 -NCPU=6 -ICPU=3
 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT
 -THREADS=1

Search information block:
Query: US-09-126-945B-2_COPY_236_243
Query length: 8
Database: N_Geneseq_0601:*
Database sequences: 730101
Database length: 313950809
Search time (sec): 104.390000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000
WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000


score_list:
Sequence            Strd Orig    ZScore    EScore Len  ! Documentation  ..
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA06617 +    8.00  166.74  0.2569    278  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF21828 +    8.00  157.74  0.8149   1087  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ50691 +    8.00  154.07  1.30     1894  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC83261 +    8.00  154.07  1.30     1894  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC83266 +    8.00  150.37  2.10     3317  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT:AAT33537 -    7.00  121.73  82.58   12412  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT97058 -    6.00  144.25  4.60       20  ! V
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA09858 -    6.00  144.25  4.60       20  ! V
/SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT:AAQ60590 -    6.00  126.55  44.50     292  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC03618 -    6.00  125.20  52.88     358  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC49470 +    6.00  124.06  61.27     426  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ80015 -    6.00  121.20  88.42     657  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF07726 +    6.00  120.63  95.10     716  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT:AAN80229 -    6.00  117.02  150.98   1236  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT:AAN82450 -    6.00  117.02  150.98   1236  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC39308 +    6.00  115.64  180.18   1523  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF14511 +    6.00  115.25  189.55   1617  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF54820 +    6.00  113.90  225.30   1983  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT:AAN70807 -    6.00  112.53  268.55   2440  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT:AAQ37484 -    6.00  112.53  268.55   2440  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT:AAN70805 -    6.00  112.18  280.98   2574  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1987.DAT:AAN70804 -    6.00  112.18  281.08   2575  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1985.DAT:AAN50014 -    6.00  112.17  281.26   2577  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT:AAQ37481 -    6.00  112.17  281.26   2577  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT:AAQ68672 +    6.00  112.17  281.26   2577  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC47006 -    6.00  112.06  285.51   2623  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA70146 -    6.00  111.64  301.10   2793  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX13293 -    6.00  110.78  336.24   3182  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT:AAV74109 +    6.00  109.27  408.37   4003  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA49442 +    6.00  102.32  994.94  11459  !
/SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF28548 -    6.00   88.28  6.0e+03 96109  !
/SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ06779 -    5.00  125.01  54.21     18  ! P
/SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX19448 +    5.00  124.31  59.26     20  ! P
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA90239 +    5.00  124.31  59.26     20  ! I
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA90202 +    5.00  122.33  76.41     27  ! I
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA90249 +    5.00  122.33  76.41     27  ! I
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA37297 +    5.00  118.27  128.74    50  ! H
/SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF54456 +    5.00  118.27  128.74    50  ! H
/SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT:AAQ69908 -    5.00  116.43  162.86    66
/SIDS1/gcgdata/geneseq/geneseqn/NA1994.DAT:AAQ69913 +    5.00  116.33  164.95    67
/SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT71227 +    5.00  116.04  171.18    70
/SIDS1/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX19447 +    5.00  116.04  171.18    70
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA90201 +    5.00  116.04  171.18    70
/SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA90238 +    5.00  116.04  171.18    70
/SIDS1/gcgdata/geneseq/geneseqn/NA1990.DAT:AAQ03206 +    5.00  112.65  264.45   117

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA06617

seq_documentation_block:
ID   AAA06617 standard; cDNA; 278 BP.
XX
AC   AAA06617;
XX
DT   13-JUN-2000  (first entry)
XX
DE   Human immunogenic prostate tumour protein cDNA sequence SEQ ID NO:398.
XX
KW   Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
KW   immunogenic; cytostatic; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200004149-A2.
XX
PD   27-JAN-2000.
XX
PF   14-JUL-1999;   99WO-US15838.
XX
PR   14-JUL-1998;   98US-0115453.
PR   14-JUL-1998;   98US-0116134.
PR   23-SEP-1998;   98US-0159812.
PR   23-SEP-1998;   98US-0159822.
PR   15-JAN-1999;   99US-0232149.
PR   15-JAN-1999;   99US-0232880.
PR   09-APR-1999;   99US-0288946.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Dillon DC,  Harlocker SL,  Yuqiu J,  Xu J,  Mitcham JL;
XX
DR   WPI; 2000-171268/15.
XX
PT   New polypeptide useful for treating and diagnosing prostate cancer
PT   comprises an immunogenic portion of prostate tumor protein -
XX
PS   Claim 50; Page 239; 263pp; English.
XX
CC   The present invention describes isolated polypeptides, comprising an
CC   immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC   and polynucleotides encoding them have cytostatic activity and can be
CC   used in vaccines and in gene therapy. The polypeptides and
CC   polynucleotides encoding them, antigen presenting cells which express
CC   the polypeptides, antibodies against the polypeptides and vaccines
CC   comprising them can be used for inhibiting the development of prostate
CC   cancer in a patient. The polypeptides can be used to generate antibodies
CC   or anti-idiotypic antibodies for passive immuno therapy. A portion of
CC   the polynucleotides encoding the polypeptides can be used as a probe or
CC   to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
CC   AAY82000 to AAY82020 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 278 BP; 56 A; 85 C; 87 G; 49 T; 1 other;


alignment_scores:
            Quality:    8.00            Length:      8
              Ratio:  1.000              Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAA06617   ..
```

```
Align seg 1/1  to: AAA06617  from: 1  to: 278

      1 ThrAspSerGluValAspSerSer 8
        ||||||||||||||||||||||||
    155 ACCGACAGCGAGGTGGACTCATCA 178

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF21828

seq_documentation_block:
ID   AAF21828 standard; DNA; 1087 BP.
XX
AC   AAF21828;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Human breast and ovarian cancer associated antigen gene SEQ ID 215.
XX
KW   Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW   nootropic; neurpprotective; antiviral; antiallergic; hepatotropic;
KW   antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW   antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW   Addison's disease; allergy; autoimmune haemolytic anaemia;
KW   autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW   multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW   cardiovascular disorder; wound healing; neurological disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO200055173-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05881.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-611515/58.
DR   P-PSDB; AAB58925.
XX
PT   New human breast and ovarian cancer associated gene sequences and the
PT   polypeptides encoded by these genes, useful in the prevention,
PT   treatment and diagnosis of cancer, immune disorders, cardiovascular
PT   disorders and neurological diseases -
XX
PS   Claim 1; Page 642-643; 1299pp; English.
XX
CC   Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC   proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC   associated with breast and ovarian cancer. Included in the invention are
CC   sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC   isolation and characterisation of the DNA and protein sequences of the
CC   invention. The breast and ovarian cancer associated DNA, protein, agonist
CC   or antagonist sequences exhibit cytostatic; immunosuppressive;
CC   nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC   antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
CC   antibacterial; antifungal; antiparasitic and cardiant activity. The
CC   polynucleotide and protein sequences are used in the diagnosis of cancer,
CC   particularly breast and ovarian cancer. The nucleic acid sequences,
CC   proteins, agonists and agonists may also be used in the diagnosis,
CC   prevention and treatment of immune disorders e.g. Addison's disease,
CC   allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC   diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC   arthritis and ulcerative colitis; cardiovascular disorders such as
CC   myocardial ischaemias; wound healing; neurological diseases such as
CC   cerebral anoxia and epilepsy; and infectious diseases.
XX
SQ   Sequence 1087 BP; 215 A; 347 C; 339 G; 182 T; 4 other;

alignment_scores:
            Quality:    8.00              Length:       8
              Ratio:   1.000                Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAF21828   ..

 Align seg 1/1  to: AAF21828  from: 1  to: 1087

       1 ThrAspSerGluValAspSerSer 8
         ||||||||||||||||||||||||
     582 ACCGACAGCGAGGTGGACTCATCA 605

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ50691

seq_documentation_block:
ID   AAZ50691 standard; cDNA; 1894 BP.
XX
AC   AAZ50691;
XX
DT   23-MAY-2000  (first entry)
XX
DE   Human prostate derived Ets factor cDNA.
XX
KW   Prostate Derived Ets factor; PDEF; human; chromosome 6p21.3; cancer;
KW   loss of heterozygosity; chromosomal translocation; linkage analysis;
KW   cytostatic; cardiant; immunosuppressive; cerebroprotective; fungicide;
KW   antibacterial; vulnerary; neuroprotective; antiparkinsonian; nootropic;
KW   anabolic; antiinflammatory; anorectic; hybridisation probe; forensic;
KW   tumour marker; diagnosis; treatment; prostate cancer; blood coagulation;
KW   autoimmune disorder; haematopoietic; immune/nervous system; stroke;
KW   neoplasm; microbial infection; tissue regeneration; heart attack;
KW   scarring; food additive; preservative; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             416..1423
FT                   /*tag= a
FT                   /product= "Human prostate derived Ets factor"
FT                   /note= "Expressed in prostate epithelium"
XX
PN   WO200006589-A1.
XX
PD   10-FEB-2000.
XX
PF   02-AUG-1999;   99WO-US17470.
XX
PR   31-JUL-1998;   98US-0126945.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
PI   Libermann TA,  Oettgen JP,  Kunsch CA,  Endress GA,  Rosen CA;
XX
DR   WPI; 2000-195255/17.
DR   P-PSDB; AAY45003.
XX
PT   Novel prostate derived polypeptide, polynucleotide useful for
PT   diagnosis, prevention and treatment of prostate cancer, autoimmune
PT   disorders, microbial infections and also as food additive or
PT   preservative  -
XX
PS   Claim 4; Fig 1; 132pp; English.
XX
CC   The present cDNA sequence encodes a human prostate derived Ets factor
CC   (PDEF), isolated from human prostate carcinoma cDNA library
CC   (ATCC No.203072). PDEF gene expression is highest in tissues involved in
CC   androgen sensitivity, with enriched expression in prostate epithelium
CC   cells. PDEF gene is mapped to the human chromosome 6p21.3
CC   region that is associated with loss of heterozygosity and chromosomal
```

```
CC   translocations in various human cancers. PDEF has cytostatic, cardiant,
CC   immunosuppressive, cerebroprotective, fungicide, antibacterial,
CC   vulnerary, neuroprotective, antiparkinsonian, nootropic, anabolic,
CC   antiinflammatory and anorectic activity. PDEF polynucleotides are useful
CC   in linkage analysis as markers, as hybridisation probes for differential
CC   identification of the tissues or cell types and as polymorphic markers
CC   for forensic purposes. PDEF is useful as prostate-specific tumour marker
CC   for the diagnosis and treatment of prostate cancer. PDEF sequences are
CC   useful for treating autoimmune disorders, haematopoietic, blood
CC   coagulation, immune and nervous system disorders, hyperproliferative
CC   disorders like, neoplasms and microbial infections, heart attacks,
CC   stroke, scarring and for tissue regeneration. They are also useful as
CC   food additives or preservatives.
XX
SQ   Sequence 1894 BP; 368 A; 653 C; 571 G; 302 T; 0 other;


alignment_scores:
            Quality:    8.00              Length:      8
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAZ50691   ..

 Align seg 1/1  to: AAZ50691  from: 1  to: 1894

       1 ThrAspSerGluValAspSerSer 8
         ||||||||||||||||||||||||
    1121 ACCGACAGCGAGGTGGACTCATCA 1144

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC83261

seq_documentation_block:
ID   AAC83261 standard; DNA; 1894 BP.
XX
AC   AAC83261;
XX
DT   16-MAR-2001  (first entry)
XX
DE   Gland-specific Ets transcription factor (GSEF) nucleotide sequence.
XX
KW   Transcription factor; gland-specific Ets transcription factor; GSEF;
KW   metastatic potential; cancer; tumour; metastasis; breast; prostate;
KW   leukaemia; lymphoma; sarcoma; melanoma; chromosome 6p21.1-6p21.3; ds.
XX
OS   Homo sapiens.
XX
PN   WO200070092-A1.
XX
PD   23-NOV-2000.
XX
PF   12-MAY-2000; 2000WO-US13173.
XX
PR   14-MAY-1999;   99US-0134112.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Kaufmann J,  Xin H,  Harrowe G;
XX
DR   WPI; 2001-041019/05.
DR   P-PSDB; AAB49628.
XX
PT   Detecting metastatic and potential metastatic cancerous cells, useful
PT   for diagnosing, prognosing, grading and staging of cancers by detecting
PT   gland-specific Ets transcription factor gene product in a biological
PT   sample from a cell
XX
PS   Disclosure; Page 81-83; 95pp; English.
XX
CC   This invention relates to a method for the detection and determination of
CC   the metastatic potential of a cell. The method comprises detecting a
CC   gland-specific Ets transcription factor (GSEF) gene product in a test
CC   sample. Detection of a GSEF gene product in the test sample in amount
CC   lower than that in a normal cell, is indicative of a cell with high
CC   metastatic potential. The method is useful for determining the metastatic
CC   potential of a cell, for the diagnosis and prognosis of cancer as well as
CC   grading and staging of cancers by detecting GSEF expression in a
CC   biological test sample. The method may also be used to monitor patients
CC   having a predisposition to develop a particular cancer. GSEF polypeptides
CC   are useful for producing antibodies, in cancer diagnosis, prognosis,
CC   grading, staging and management of breast and prostate tumours, and in
CC   detecting polymorphisms in the sequence. GSEF genes and proteins are also
CC   useful in gene therapy. GSEF gene product expression levels can be used
CC   in conjunction with any tissue in which an alteration in GSEF gene
CC   product expression levels is associated with development of a
CC   cancer-associated phenotype. Cancers, which can be monitored include
CC   cancers of the prostate, cervix, lung and colon, melanomas, colorectal
CC   adenocarcinomas, Wilm's tumour, retinoblastoma, sarcomas, myosarcomas,
CC   lung carcinomas, leukaemia, and lymphomas. The GSEF gene is located on
CC   human chromosome 6, specifically at 6p21.1-6p21.3. The present sequence
CC   represents the DNA encoding GSEF.
XX
SQ   Sequence 1894 BP; 368 A; 653 C; 571 G; 302 T; 0 other;


alignment_scores:
            Quality:    8.00              Length:      8
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAC83261   ..

 Align seg 1/1  to: AAC83261  from: 1  to: 1894

       1 ThrAspSerGluValAspSerSer 8
         ||||||||||||||||||||||||
    1121 ACCGACAGCGAGGTGGACTCATCA 1144

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC83266

seq_documentation_block:
ID   AAC83266 standard; DNA; 3317 BP.
XX
AC   AAC83266;
XX
DT   16-MAR-2001  (first entry)
XX
DE   Gland-specific Ets transcription factor (GSEF) cDNA sequence.
XX
KW   Transcription factor; gland-specific Ets transcription factor; GSEF;
KW   metastatic potential; cancer; tumour; metastasis; breast; prostate;
KW   leukaemia; lymphoma; sarcoma; melanoma; chromosome 6p21.1-6p21.3; ss.
XX
OS   Homo sapiens.
XX
PN   WO200070092-A1.
XX
PD   23-NOV-2000.
XX
PF   12-MAY-2000; 2000WO-US13173.
XX
PR   14-MAY-1999;   99US-0134112.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Kaufmann J,  Xin H,  Harrowe G;
XX
DR   WPI; 2001-041019/05.
DR   P-PSDB; AAB49633.
XX
PT   Detecting metastatic and potential metastatic cancerous cells, useful
PT   for diagnosing, prognosing, grading and staging of cancers by detecting
PT   gland-specific Ets transcription factor gene product in a biological
PT   sample from a cell
```

```
XX
PS   Disclosure; Fig 2; 95pp; English.
XX
CC   This invention relates to a method for the detection and determination of
CC   the metastatic potential of a cell. The method comprises detecting a
CC   gland-specific Ets transcription factor (GSEF) gene product in a test
CC   sample. Detection of a GSEF gene product in the test sample in amount
CC   lower than that in a normal cell, is indicative of a cell with high
CC   metastatic potential. The method is useful for determining the metastatic
CC   potential of a cell, for the diagnosis and prognosis of cancer as well as
CC   grading and staging of cancers by detecting GSEF expression in a
CC   biological test sample. The method may also be used to monitor patients
CC   having a predisposition to develop a particular cancer. GSEF polypeptides
CC   are useful for producing antibodies, in cancer diagnosis, prognosis,
CC   grading, staging and management of breast and prostate tumours, and in
CC   detecting polymorphisms in the sequence. GSEF genes and proteins are also
CC   useful in gene therapy. GSEF gene product expression levels can be used
CC   in conjunction with any tissue in which an alteration in GSEF gene
CC   product expression levels is associated with development of a
CC   cancer-associated phenotype. Cancers, which can be monitored include
CC   cancers of the prostate, cervix, lung and colon, melanomas, colorectal
CC   adenocarcinomas, Wilm's tumour, retinoblastoma, sarcomas, myosarcomas,
CC   lung carcinomas, leukaemia, and lymphomas. The GSEF gene is located on
CC   human chromosome 6, specifically at 6p21.1-6p21.3. The present sequence
CC   represents the cDNA encoding GSEF.
XX
SQ   Sequence 3317 BP; 710 A; 1026 C; 970 G; 611 T; 0 other;


alignment_scores:
            Quality:    8.00              Length:      8
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAC83266   ..

 Align seg 1/1  to: AAC83266  from: 1  to: 3317

       1 ThrAspSerGluValAspSerSer 8
         ||||||||||||||||||||||||
    2520 ACCGACAGCGAGGTGGACTCATCA 2543

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA1996.DAT:AAT33537

seq_documentation_block:
ID   AAT33537 standard; DNA; 12412 BP.
XX
AC   AAT33537;
XX
DT   15-FEB-1998  (first entry)
XX
DE   BCG deletion region 3 and flanking sequences.
XX
KW   BCG delta 3; virulence; avirulence; attenuation; gene deletion;
KW   mycobacteria; vaccine; infection; marker; ss.
XX
OS   Mycobacterium bovis strain BCG.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1406..10673
FT                   /*tag= a
FT                   /note= "BCG delta 1 deletion region"
XX
PN   WO9625519-A1.
XX
PD   22-AUG-1996.
XX
PF   15-FEB-1996;   96WO-US01938.
XX
PR   17-FEB-1995;   95US-0390878.
XX
PA   (PATH-) PATHOGENESIS CORP.
XX
PI   Mahairas GG,  Stover CK;
XX
DR   WPI; 1996-393419/39.
XX
PT   Detecting markers for avirulence in Mycobacterium - used in
PT   production of vaccines against bacterial infection, and to detect
PT   bacterial infection
XX
PS   Example 1; Fig 3; 66pp; English.
XX
CC   This DNA sequence comprises Mycobacterium bovis BCG deletion
CC   sequence BCGdelta3.  A specific genetic deletion of this region
CC   results in an avirulence phenotype of the mycobacterium.  2 Other
CC   deletion regions (see AAT33535 and AAT33536) have also been detected.
CC   Identification involved screening a BCG cosmid library with a
CC   radiolabeled probe obtained following DNA subtraction between
CC   virulent Mycobacterium tuberculosis H37Rv and avirulent BCG.
CC   The deletions provide useful markers for the identification of an
CC   avirulent, or a virulent, mycobacterial phenotype.  Determination
CC   of avirulence requires the detection of the presence or absence of
CC   the deletion; the deletions are detected either by detecting the
CC   presence or absense of deletion junctions (see AAT33538-46), or by
CC   detecting the presence or absence of the sequences contained within
CC   the deletion.  Deletion polypeptides are used as components of
CC   immunological assays and in vaccines.
XX
SQ   Sequence 12412 BP; 2146 A; 4273 C; 3876 G; 1917 T; 200 other;


alignment_scores:
            Quality:    7.00              Length:      7
              Ratio:  1.000                Gaps:      0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAT33537/rev   ..

 Align seg 1/1  to reverse of: AAT33537  from: 1  to: 12412

       2 AspSerGluValAspSerSer 8
         |||||||||||||||||||||
    1459 GATTCGGAGGTAGACAGCAGC 1439

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT97058

seq_documentation_block:
ID   AAT97058 standard; DNA; 20 BP.
XX
AC   AAT97058;
XX
DT   20-MAR-1998  (first entry)
XX
DE   VEE target specific primer 1569.
XX
KW   PCR primer; amplify; Venezuelan equine encephalitis virus; VEE;
KW   nucleic acid detection; replication reaction; clinical sample analysis;
KW   food analysis; crop analysis; homogeneous detection probe system; ss.
XX
OS   Synthetic.
OS   Venezuelan equine encephalitis virus.
XX
PN   WO9732044-A1.
XX
PD   04-SEP-1997.
XX
PF   27-FEB-1997;   97WO-US02892.
XX
PR   01-MAR-1996;   96US-0012636.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Ebersole RC,  Fitzpatrick-Mcelligott S,  Hendrickson ER;
```

```
PI   Perry MP;
XX
DR   WPI; 1997-448704/41.
XX
PT   Detecting replicated target nucleic acid - by incorporating a label
PT   in presence of a unreplicable, labelled detection probe, used for
PT   analysis of clinical samples, food, etc
XX
PS   Example 1; Page 30; 69pp; English.
XX
CC   This sequence represents a primer for a Venezuelan equine encephalitis
CC   virus (VEE) DNA sequence, and can be used in the method of the invention.
CC   The method of the invention is for detecting a target nucleic acid
CC   sequence (I) in a replication reaction. The method comprises:
CC   (i) replicating at least one (I) with a composition containing first
CC   label (L1) that can be incorporated into replicated nucleic acid and a
CC   detection probe (DP) consisting of a second label (L2), target domain and
CC   replication inhibiting domain that prevents the DP from participating in
CC   the replication reaction, to produce a hybrid of DP and at least one
CC   replicated (I); (ii) immobilising the hybrid through L1 or L2; and (iii)
CC   detecting presence of immobilised hybrid. The method is used with any
CC   replication process, including PCR in cells, e.g. for analysis of
CC   clinical samples, foods, crops, soil, water etc. The method using a
CC   homogeneous detection probe system allows real-time monitoring of product
CC   formation and is not restricted by base composition of probes or primers,
CC   melting or annealing temperatures or PCR cycling conditions. The presence
CC   of probes throughout the reaction does not inhibit replication or
CC   decrease yield, and eliminates the need for an additional hybridisation
CC   stage, resulting in faster and more efficient assay. Also detection of
CC   unwanted sequences is reduced, electrophoresis is not required and
CC   products can be detected in several different formats.
XX
SQ   Sequence 20 BP; 4 A; 6 C; 4 G; 6 T; 0 other;


alignment_scores:
            Quality:    6.00              Length:      6
              Ratio:  1.000                Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAT97058/rev   ..

 Align seg 1/1  to reverse of: AAT97058  from: 1  to: 20

       3 SerGluValAspSerSer 8
         ||||||||||||||||||
      19 TCAGAAGTGGACAGCAGT 2

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA09858

seq_documentation_block:
ID   AAA09858 standard; DNA; 20 BP.
XX
AC   AAA09858;
XX
DT   05-JUL-2000  (first entry)
XX
DE   Venezuelan equine encephalitis virus gene primer VE1569.
XX
KW   Detection; non-denatured nucleic acid; test strip; capture zone; dnaK;
KW   blocking agent; nonionic surfactant; chromatography; capillary migration;
KW   human; veterinary medicine; agricultural science; food science; ss;
KW   bacteria; virus; antibiotic resistance; malignant cell; PCR primer.
XX
OS   Venezuelan equine encephalitis virus.
XX
PN   US6037127-A.
XX
PD   14-MAR-2000.
XX
PF   26-NOV-1997;   97US-0979269.
XX
PR   31-MAR-1994;   94US-0221769.
PR   20-SEP-1995;   95US-0530795.
PR   27-MAY-1997;   97US-0863265.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Ebersole RC,  Payne MS,  Fitzpatrick-McElligott S,  Majarian WR;
PI   Rafalski JA,  Hendrickson ER;
XX
DR   WPI; 2000-269884/23.
XX
PT   Detecting non-denatured nucleic acid useful for medical diagnosis
PT   comprises specific capture on test strip and detection of incorporated
PT   or associated reporter -
XX
PS   Example 9; Column 36; 39pp; English.
XX
CC   The invention relates to methods of detecting the presence of
CC   non-denatured nucleic acid (I) in a buffered test sample by applying
CC   the sample to a test strip that has, in a capture zone, an immobilized
CC   single-stranded nucleic acid (II) complementary to part of (I). After
CC   incubation, the strip is washed and treated with a signal-generating
CC   substance (A) that reacts with a reporter conjugate located within, or
CC   bound to, (I). The sample includes a blocking agent and nonionic
CC   surfactant and is applied to one end of the strip made of
CC   chromatographic material that transports sample laterally by capillary
CC   migration. The strip is incubated for at least 5 mins, then washed to
CC   remove unbound nucleic acid from the capture zone. The signal from the
CC   capture zone is then compared with that from some other part of the
CC   strip. The method is used to identify (I) in human or veterinary
CC   medicine, or in agricultural and food science, particularly to detect or
CC   identify bacteria and viruses, to screen microbes for antibiotic
CC   resistance, and to detect malignant cells. Primers AAA09856-A098590 were
CC   used to PCR amplify fragments of the Venezuelan equine encephalitis virus
CC   genome which used as an example of a target sequence to be detected by
CC   the method of the invention.
XX
SQ   Sequence 20 BP; 4 A; 6 C; 4 G; 6 T; 0 other;


alignment_scores:
            Quality:    6.00              Length:      6
              Ratio:  1.000                Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAA09858/rev   ..

 Align seg 1/1  to reverse of: AAA09858  from: 1  to: 20

       3 SerGluValAspSerSer 8
         ||||||||||||||||||
      19 TCAGAAGTGGACAGCAGT 2

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA1993.DAT:AAQ60590

seq_documentation_block:
ID   AAQ60590 standard; cDNA; 292 BP.
XX
AC   AAQ60590;
XX
DT   16-MAR-1994  (first entry)
XX
DE   Human brain Expressed Sequence Tag EST02601.
XX
KW   Gene transcription product; genetic markers; tagging; in vivo;
KW   transcription; mapping; locations; chromosomes; chromosomal; ss.
XX
OS   Homo sapiens.
XX
PN   WO9316178-A.
XX
PD   19-AUG-1993.
```

```
XX
PF   12-FEB-1993;   93WO-US01294.
XX
PR   12-FEB-1992;   92US-0837195.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
PI   Adams MD,  Moreno RF,  Venter CJ;
XX
DR   WPI; 1993-272882/34.
XX
PT   Enriched oligonucleotides and corresp. sequences - used as
PT   markers for human genes transcribed in-vivo, facilitate tagging
PT   of most human genes
XX
PS   Example 4; Page 358; 500pp; English.
XX
CC   The Expressed Sequence Tag was isolated from a human brain cDNA
CC   library as part of a large set of ESTs which can be used as markers
CC   for human genes transcribed in vivo. They can be used to facilitate
CC   tagging of most human genes, for mapping locations of expressed genes
CC   on chromosomes, for individual or forensic identification, for mapping
CC   locations of disease-associated genes, for identification of tissue
CC   type, and for prepn. of antisense sequences, probes and constructs.
CC   EST02601 has an "excellent" coding probability as evaluated using the
CC   coding-region prediction program CRM. See also AAQ59041-Q61440.
XX
SQ   Sequence 292 BP; 81 A; 80 C; 57 G; 69 T; 5 other;


alignment_scores:
            Quality:    6.00              Length:.      6
              Ratio:  1.000                Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAQ60590/rev    ..

 Align seg 1/1  to reverse of: AAQ60590  from: 1  to: 292

       3 SerGluValAspSerSer 8
         ||||||||||||||||||
     271 AGTGAAGTTGACAGCAGT 254

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC03618

seq_documentation_block:
ID   AAC03618 standard; cDNA; 358 BP.
XX
AC   AAC03618;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 3616.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   P-PSDB; AAG03612.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 1; SEQ ID 3616; 71pp + CD-ROM; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. An ORF has been identified within the
CC   sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC   derived from 30 different tissues. EST sequences usually correspond
CC   mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC   often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC   well suited for isolating cDNA sequences derived from the 5' ends of
CC   mRNAs and even in those cases where longer cDNA sequences have been
CC   obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC   mRNAs with intact 5' ends and can therefore be used to obtain full length
CC   cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC   gene therapy and chromosome mapping procedures. They are used to obtain
CC   upstream regulatory sequences and to design expression and secretion
CC   vectors.
XX
SQ   Sequence 358 BP; 107 A; 69 C; 71 G; 109 T; 2 other;


alignment_scores:
            Quality:    6.00              Length:       6
              Ratio:  1.000                Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAC03618/rev    ..

 Align seg 1/1  to reverse of: AAC03618  from: 1  to: 358

       3 SerGluValAspSerSer 8
         ||||||||||||||||||
     349 AGTGAAGTGGATTCCTCC 332

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC49470

seq_documentation_block:
ID   AAC49470 standard; DNA; 426 BP.
XX
AC   AAC49470;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 61280.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
```

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

```
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

alignment_scores:
            Quality:    6.00              Length:      6
              Ratio:  1.000                Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAC49470    ..

 Align seg 1/1  to: AAC49470  from: 1  to: 426

       1 ThrAspSerGluValAsp 6
         ||||||||||||||||||
     322 ACAGATTCTGAAGTTGAT 339

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ80015

seq_documentation_block:
ID   AAZ80015 standard; cDNA; 657 BP.
XX
AC   AAZ80015;
XX
DT   07-APR-2000  (first entry)
XX
DE   Human colon cancer cell line SW480 cDNA clone SEQ ID NO:99.
XX
KW   Human; gene expression product; diagnosis; tumour; colon cancer;
KW   colorectal adenocarcinoma; cell line SW480; cell proliferation;
KW   cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
KW   hyperplasia; ds.
XX
OS   Homo sapiens.
XX
PN   WO9964576-A2.
XX
PD   16-DEC-1999.
XX
PF   09-JUN-1999;   99WO-IB01062.
XX
PR   10-JUN-1998;   98US-0088801.
XX
PA   (FARB ) BAYER CORP.
XX
PI   Endege WO,  Steinmann KE,  Astle JH,  Burgess CC,  Bushnell SE;
PI   Carroll E,  Catino TJ,  Derti A,  Ford DM,  Lewis ME,  Monahan JE;
PI   Schlegel R;
XX
DR   WPI; 2000-087220/07.
XX
PT   Novel nucleic acids, used to develop products for the diagnosis and
PT   treatment of disorders involving unwanted cell proliferation,
PT   particularly cancers, especially colon cancer  -
XX
PS   Claim 1; Page 171; 469pp; English.
XX
CC   AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from
CC   the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
CC   cDNA clones can be used to generate antisense oligonucleotides which
CC   can be used for antisense therapy. Methods and products from the present
CC   invention can be used for identifying and/or classifying cancerous cells
CC   present in a human tumour, particularly in solid tumours, e.g.
CC   carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
CC   can be used for developing agents for the diagnosis and treatment of
CC   disorders involving unwanted cell proliferation, such as neoplasia,
CC   dysplasia or hyperplasia.
XX
SQ   Sequence 657 BP; 137 A; 146 C; 153 G; 203 T; 18 other;


alignment_scores:
            Quality:    6.00              Length:      6
              Ratio:  1.000                Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAZ80015/rev    ..

 Align seg 1/1  to reverse of: AAZ80015  from: 1  to: 657

       3 SerGluValAspSerSer 8
         ||||||||||||||||||
     199 TCAGAGGTGGACAGCTCA 182

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF07726

seq_documentation_block:
ID   AAF07726 standard; cDNA; 716 BP.
XX
AC   AAF07726;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Fusarium venenatum EST SEQ ID NO:249.
XX
KW   Multiple gene expression; filamentous fungal cell; EST;
KW   expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW   Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW   culture condition; environmental stress; spore morphogenesis;
KW   metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS   Fusarium venenatum.
XX
PN   WO200056762-A2.
XX
PD   28-SEP-2000.
XX
PF   22-MAR-2000; 2000WO-US07781.
XX
PR   22-MAR-1999;   99US-0273623.
XX
PA   (NOVO ) NOVO NORDISK BIOTECH INC.
PA   (NOVO ) NOVO NORDISK AS.
XX
PI   Berka RM,  Rey MW,  Shuster JR,  Kauppinen S,  Clausen IG,  Olsen PB;
XX
DR   WPI; 2000-594572/56.
XX
PT   Monitoring differential expression of genes in filamentous fungal cells
PT   uses fluorescence-labeled nucleic acids isolated from the cells and a
PT   substrate of expressed sequence tags -
XX
PS   Claim 86; Page 478; 3161pp; English.
```

```
XX
CC   The present invention describes a method for monitoring differential
CC   expression of genes in a first filamentous fungal (FF) cell relative to
CC   expression of the same genes in one or more second filamentous fungal
CC   cells. The method uses fluorescence-labeled nucleic acids isolated from
CC   the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC   are used in the methods for monitoring differential expression of genes
CC   in a first filamentous fungal (FF) cell relative to expression of the
CC   same genes in one or more second filamentous fungal cells. Monitoring
CC   the global expression of genes from FF cells allows the production
CC   potential of the microorganisms to be improved. New genes may be
CC   discovered, possible functions of unknown open reading frames can be
CC   identified and gene copy number variation and stability can be
CC   monitored. The expression of genes can be used to study how FF cells
CC   adapt to changes in culture conditions, environmental stress, spore
CC   morphogenesis, recombination, metabolic or catabolic pathway
CC   engineering. Using ESTs provides several advantages over genomic or
CC   random cDNA clones including elimination of redundancy as one spot on an
CC   array equals one gene or open reading frame, and organisation of the
CC   microarrays based on function of the gene products to facilitate
CC   analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC   Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
CC   niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
CC   AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
CC   all specifically claimed in the present invention.
XX
SQ   Sequence 716 BP; 172 A; 214 C; 172 G; 157 T; 1 other;


alignment_scores:
           Quality:    6.00              Length:      6
             Ratio:  1.000                Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAF07726   ..

 Align seg 1/1  to: AAF07726  from: 1  to: 716

       1 ThrAspSerGluValAsp 6
         ||||||||||||||||||
     612 ACCGACAGCGAAGTCGAC 629

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT:AAN80229

seq_documentation_block:
ID   AAN80229 standard; DNA; 1236 BP.
XX
AC   AAN80229;
XX
DT   03-DEC-1990  (first entry)
XX
DE   Bioadhesive precursor protein gene cDNA 52.
XX
KW   Bioadhesive precursor protein; cDNA 52; mussels; barnacles; oysters;
KW   biomedical adhesive; sealants; wound healing; ss.
XX
OS   Mytilus edulis.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..1004
FT                   /*tag=  a
FT                   /product=bioadhesive precursor protein
XX
PN   WO8803953-A.
XX
PD   02-JUN-1988.
XX
PF   24-NOV-1987;   87WO-US03048.
XX
PR   07-AUG-1987;   87US-0082456.
XX
PA   (GENE-) GENEX CORP.
XX
PI   Maugh KJ,  Anderson DM,  Strausberg R,  Strausberg SL,  McCandliss R;
XX
DR   WPI; 1988-161622/23.
DR   P-PSDB; AAP82971.
XX
PT   DNA coding for bio -adhesive precursor protein - obtd. from mussels,
PT   barnacles and oysters and used esp. to prepare biomedical adhesive.
XX
PS   Disclosure; ; pp; English.
XX
CC   mRNA from the phenol gland of the foot of M. edulis is isolated for
CC   the synthesis of cDNA.  The cDNA is inserted into a bacteriophage
CC   and E. coli transfected and cultured.  Five clones were isolated and
CC   characterised, one of which is clone 52.  The sequence has been
CC   divided to show 26 subunits, including 23 decapeptides and 3
CC   hexapeptides, and an amino-terminal proline-rich segment.
CC   There is also a direct repeat of DNA sequences in clone 52.
CC   This cDNA clone and two of the other (55 (AAN80230) and 56 (AAN80231))
CC   contain some momcoding DNA sequence (not shown) which is believed
CC   to be derived from intron sequences.
CC   Clone 14-1 (AAN80228) and 52 display extensive homology at their 3'
CC   ends. In particular, the last 138 codons of the translated regions are
CC   identical and include codons for a hexapeptide followed by 12
CC   decapeptides (beginning at codon 75 of clone 14-1 and codon 205
CC   of clone 52.  The DNA segment contg. the coding regions of
CC   bioadhesive precursor protein of clone 14-1 and 52 was obtained from
CC   lambda gt10 cloning vectors by EcoRI digestion and subcloned into the
CC   EcoRI site of plasmid pGX2627.
CC   The protein is chemically or enzymatically hydroxylated to produce
CC   adhesive compsns. having excellent properties in wet environments. The
CC   product is partic. useful as a biomedical adhesive or sealant e. g. in
CC   wound healing or as a marine or dental adhesive.  It can be used as a
CC   thin film membrane.
CC   See also AAN80228-34.
XX
SQ   Sequence 1236 BP; 464 A; 338 C; 125 G; 309 T; 0 other;


alignment_scores:
           Quality:    6.00              Length:      6
             Ratio:  1.000                Gaps:       0
 Percent Similarity: 100.000   Percent Identity: 100.000

alignment_block:
 US-09-126-945B-2_COPY_236_243 x AAN80229/rev   ..

 Align seg 1/1  to reverse of: AAN80229  from: 1  to: 1236

       2 AspSerGluValAspSer 7
         |||||||||||||||||
     204 GACTCGGAGGTGGACTCG 187

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA1988.DAT:AAN82450

seq_documentation_block:
ID   AAN82450 standard; cDNA; 1236 BP.
XX
AC   AAN82450;
XX
DT   06-MAR-1992  (first entry)
XX
DE   Sequence for a gene identified as cDNA clone 52 which codes for
DE   a bioadhesive precursor protein.
XX
KW   Bioadhesive; wound healing; bonding; recombinant adhesive; ss.
XX
OS   Mytilus edulis.
XX
FH   Key             Location/Qualifiers
FT   CDS             3..1022
FT                   /*tag=  a
XX
```

PN WO8807076-A.
XX
PD 22-SEP-1988.
XX
PF 11-MAR-1988; 88WO-US00876.
XX
PR 12-MAR-1987; 87US-0025243.
XX
PA (GENE-) GENEX CORP.
XX
PI Maugh KJ, Anderson DM, Strausberg SL, Strausberg R;
PI Wei T;
XX
DR WPI; 1988-285533/40.
DR P-PSDB; AAP83194.
XX
PT Bio-adhesive precursor protein analogues prodn. - by DNA
PT recombinant techniques, subsequent hydroxylation, gives an
PT adhesive for use in wet environment
XX
PS Example; Fig 14; 101pp; English.
XX
CC The bioadhesive precursor protein analogue of the invention can be
CC hydroxylated (prod. claimed) and used as an adhesive in wet
CC environments. The resulting bioadhesives have marine, biomedical and
CC dental applications. They may be used e.g. in wound healing in the
CC same manner as fibrin. While the decapeptide ala-lys-pro-ser-tyr-pro-
CC pro-thr-tyr-lys is repeated many times in the polyphenolic adhesive
CC protein of M.edulis, the examination of cDNA sequences encoding
CC portions of this protein (see U.S. Patent Application Serial No. 933,
CC 945) reveals that many other repetitive decapeptide and hexapeptide
CC sequences are also present in the protein, and that these other
CC sequences may constitute the majority of AA sequence of the
CC polyphenolic adhesive protein of M.edulis. For example, in cDNA
CC clone 14-1 nineteen decapeptides and one hexapeptide are encoded.
XX
SQ Sequence 1236 BP; 462 A; 340 C; 124 G; 310 T; 0 other;

alignment_scores:
Quality: 6.00 Length: 6
Ratio: 1.000 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-126-945B-2_COPY_236_243 x AAN82450/rev ..

Align seg 1/1 to reverse of: AAN82450 from: 1 to: 1236

```
       2 AspSerGluValAspSer 7
         |||||||||||||||||
     204 GACTCGGAGGTGGACTCG 187
```